



King's Research Portal

DOI:

[10.1016/j.jprot.2017.04.019](https://doi.org/10.1016/j.jprot.2017.04.019)

Document Version

Version created as part of publication process; publisher's layout; not normally made publicly available

[Link to publication record in King's Research Portal](#)

Citation for published version (APA):

Politis, A., & Schmidt, C. (2017). Structural characterisation of medically relevant protein assemblies by integrating mass spectrometry with computational modelling. *Journal Of Proteomics*.
<https://doi.org/10.1016/j.jprot.2017.04.019>

Citing this paper

Please note that where the full-text provided on King's Research Portal is the Author Accepted Manuscript or Post-Print version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version for pagination, volume/issue, and date of publication details. And where the final published version is provided on the Research Portal, if citing you are again advised to check the publisher's website for any subsequent corrections.

General rights

Copyright and moral rights for the publications made accessible in the Research Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognize and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the Research Portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the Research Portal

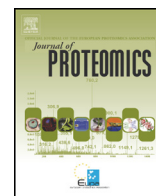
Take down policy

If you believe that this document breaches copyright please contact librarypure@kcl.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.



Contents lists available at ScienceDirect

Journal of Proteomics

journal homepage: www.elsevier.com/locate/jprot

Review

Structural characterisation of medically relevant protein assemblies by integrating mass spectrometry with computational modelling

Argyris Politis^{a,*}, Carla Schmidt^{b,*}^a Department of Chemistry, Kings College London, London, United Kingdom^b Interdisciplinary Research Center HALOmem, Martin Luther University Halle-Wittenberg, Halle/Saale, Germany

ARTICLE INFO

Article history:

Received 20 December 2016

Received in revised form 13 April 2017

Accepted 18 April 2017

Available online xxxx

Keywords:

Proteomics

Structural mass spectrometry

Computational modelling

Integrative modelling

ABSTRACT

Structural mass spectrometry with its various techniques is a powerful tool for the structural elucidation of medically relevant protein assemblies. It delivers information on the composition, stoichiometries, interactions and topologies of these assemblies. Most importantly it can deal with heterogeneous mixtures and assemblies which makes it universal among the conventional structural techniques. In this review we summarise recent advances and challenges in structural mass spectrometric techniques. We describe how the combination of the different mass spectrometry-based methods with computational strategies enable structural models at molecular levels of resolution. These models hold significant potential for helping us in characterizing the function of protein assemblies related to human health and disease.

Significance: In this review we summarise the techniques of structural mass spectrometry often applied when studying protein-ligand complexes. We exemplify these techniques through recent examples from literature that helped in the understanding of medically relevant protein assemblies. We further provide a detailed introduction into various computational approaches that can be integrated with these mass spectrometric techniques. Last but not least we discuss case studies that integrated mass spectrometry and computational modelling approaches and yielded models of medically important protein assembly states such as fibrils and amyloids.

© 2017 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Contents

1. Introduction	0
2. Structural mass spectrometry	0
2.1. Proteomics reveals changes in protein expression and post-translational modifications	0
2.2. Cross-linking to study binary protein interactions	0
2.3. Hydrogen-deuterium-exchange to study ligand binding and dynamics	0
2.4. Native ion mobility-mass spectrometry reveals shape and topology of protein assemblies	0
3. Computational modelling techniques	0
3.1. Restraint-guided computational modelling	0
3.2. Ensemble analysis	0
3.3. Scoring structural models against multiple experimental data	0
3.4. Software tools for integrative modelling	0
3.5. Modelling protein assembly dynamics	0
4. Integrating structural mass spectrometry and computational modelling - application examples.	0
4.1. Early assembly mechanism for amyloidogenic β 2-microglobulin revealed by cross-linking and native mass spectrometry	0
4.2. Insulin fibrillation examined by ion mobility mass spectrometry and molecular modelling	0
5. Conclusions	0
Conflict of interest	0
Acknowledgements	0
References	0

* Corresponding author.

E-mail addresses: argyris.politis@kcl.ac.uk (A. Politis), carla.schmidt@biochemtech.uni-halle.de (C. Schmidt).

1. Introduction

Proteins are key players in life executing critical tasks for cellular functions. These functions are often linked with their structural arrangements as well as their dynamic interactions with ligands such as proteins/peptides, nucleotides, carbohydrates or lipids. The structural analysis of proteins and the complexes they form with their interaction partners is therefore of paramount importance. Recent improvements in mass spectrometry made it an ideal tool for the structural analysis of proteins and their ligands [1–5]. It delivers information on the composition, stoichiometries, protein interactions and interaction networks as well as ligand binding [6]. There is a variety of techniques linked with mass spectrometry including proteomics, (chemical or UV) cross-linking, labelling techniques (such as hydrogen/deuterium-exchange) or mass spectrometry of intact protein assemblies, the latter is often coupled with ion mobility (IM) spectroscopy (see Table 1) [6].

Mass spectrometric techniques target either the peptides produced by hydrolysis of the proteins using proteases (i.e. bottom-up approaches) or the intact proteins. The latter distinguishes between top-down proteomics, i.e. the fragmentation of the intact proteins during analysis, or the analysis of the intact proteins using “native” mass spectrometry [7]. The advantage of bottom-up approaches is the opportunity to unambiguously identify specific sites such as modified amino acid residues (proteomics) or binary interaction sites (cross-linking). A top-down approach, on the other hand, allows the quantitative determination of these sites in the same experiment. Native mass spectrometry determines protein complex’ stoichiometries along with protein interactions and specificity of the ligands.

As the structure of a protein-ligand complex is directly linked with function, its malfunction is just as dependent. Nowadays mass spectrometry is playing an important role in structure elucidation of pharmaceutically relevant proteins as well as drug binding [8]. However, the information obtained from the various mass spectrometric techniques often yields low-resolution structures. One way to increase the degree of resolution and gain information unattainable by a single method is to bring together the information derived from various techniques [9,10]. This has prompted researchers to develop computational tools and methods that allowed encoding structural information derived from various mass spectrometry-based experiments (as well as other techniques) into diverse sets of modelling restraints, i.e. establishing integrative modelling [11,12]. The integration of these restraints into sophisticated algorithms for 3D model generation enabled

determining the structure and in many cases the conformational dynamics of proteins and their complexes [11]. The wealth of computational tools and algorithms for structure-based prediction are exemplified by homology modelling [13], and docking strategies [14], for instance in drug discovery [15] or techniques such as molecular dynamics simulations [16] and de novo modelling.

In this review we summarise the advances of structural mass spectrometry in the study of protein-ligand complexes. We exemplify the different techniques through recent examples from literature that helped in the understanding of medically relevant protein assemblies. We further provide a detailed introduction into various computational approaches that can be used to bring together the information from the various mass spectrometric techniques. Last but not least we discuss case studies that integrated mass spectrometry and computational modelling approaches and yielded models of medically important protein assembly states such as fibrils and amyloids.

2. Structural mass spectrometry

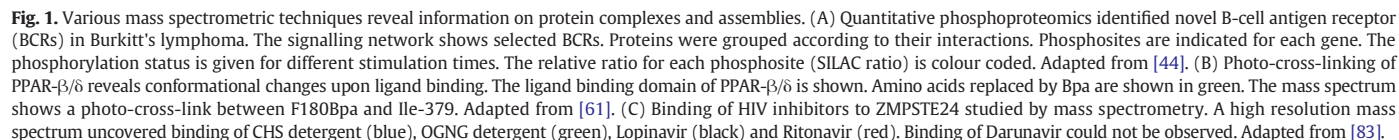
2.1. Proteomics reveals changes in protein expression and post-translational modifications

The proteome is the entire set of proteins in a cell, an organism or a tissue at a defined time point and under defined conditions [17]. Proteomics, consequently, is the large-scale study of the proteome and nowadays mostly involves mass spectrometry-based protein identification [18–21]. For this, proteins are usually hydrolysed using specific endoproteases. The specific peptide mixture is then analysed in a mass spectrometer. Peptide and fragment masses lead to the identification of the protein(s) after database searching [22]. Depending on the sample complexity separation techniques on the peptide or protein level are included in the proteomics workflow [23–25]. Recent improvements in mass spectrometry enabled the identification of post-translational modifications [26,27] and relative or absolute quantification of proteins [28,29] on a routine basis. Notably, entire proteomes can nowadays be explored in very short time spans [30,31]. As such several proteomes and their interactomes have been described to a high standard. Examples include the studies on *Saccharomyces cerevisiae* [32–34], *Drosophila melanogaster* [35,36] and even humans [37–39].

Of the available techniques, quantitative mass spectrometry plays a major role in clinical proteomics [40–42]. In particular, the quantification of post-translational modifications, which are the regulatory

Table 1
Overview on MS techniques discussed in this review. The principle, experimental workflows, commonly used instrumentation, limitations as well as the outcome are given for each technique.

MS technique	Principle	Experimental procedure	Instrumentation	Limitations	Outcome
Proteomics	Identification of proteins by fragmentation/sequencing of peptides	(i) Digestion of proteins, (ii) LC-MS/MS of peptides, (iii) database searching	High speed, high sensitivity mass spectrometers (mostly Orbitrap, Q-ToF or Q-Trap instruments)	Protein interactions only identified indirectly	Identification of proteins and post-translational modifications; quantification; interactomes
Cross-linking	Identification of binary protein interactions by fragmentation/sequencing of cross-linked di-peptides	(i) Cross-linking of proteins, (ii) digestion of proteins, (iii) LC-MS/MS of peptides, (iv) identification of cross-linked di-peptides by specialised software; (*) optional pre-fractionation of proteins or peptides	High sensitivity, high resolution mass spectrometers (mostly Orbitrap instruments)	Only binary interactions; cross-linking yield dependent on protein sequence and cross-linking chemistry	Identification of protein interactions; protein networks; distance restraints
H/D-X	Exchange of backbone protons by deuterium to study solvent accessibility	(i) Exchange of protons by deuterium, (ii) analysis of intact proteins or peptides after digestion, (iii) increase/decrease of protein/peptide masses, (iv) structural analysis of changes in H/D-X (peptides)	Mostly Q-ToF instruments; in some cases automated procedure using pipetting robots	Back exchange of deuterium; structural information of protein required; data analysis difficult	Solvent accessibility; structural changes/dynamics; ligand binding
Native ion-mobility MS	Transfer of intact protein assemblies into the gas phase of a mass spectrometer	(i) Exchanging purification buffer to aqueous, volatile buffer, (ii) manual MS analysis of protein complexes, (iii) manual data analysis (using specialised software)	Modified instruments, usually Q-ToF mass spectrometers (in some cases modified Orbitrap instruments)	Protein complexes need to be kept intact; modified instruments required; no direct interaction sites	Identification of protein stoichiometries and interactions; topology; ligand binding; collision cross-sections; shape restraints



A powerful way to study ligand binding to proteins classified as drug targets is by monitoring the exchange of labile protons for bulk deuterium in solution. This is commonly used in hydrogen-deuterium-exchange (H/D-X) where amide protons of the protein backbones are exchanged by deuterium in aqueous D₂O-containing buffers [62,63]. The uptake of deuterium depends on temperature and pH as well as solvent accessibility and hydrogen bonding of the amide protons; the latter two are indicators of protein conformational changes. Global analysis of

the intact proteins usually involves differential comparison of protein conformations (e.g. bound versus unbound states). Digestion of the labelled proteins provides structural insights into particular sites of the proteins and their ability to exchange protons. As the exchange reaction is usually quenched at low pH (<2.5) pepsin is the enzyme of choice [62, 63]. Comparing the time-dependent deuterium uptake of a protein under different condition gives rise to changes in solvent accessibility for example upon folding of a protein [64] or binding of a ligand [65]. Encoding such information into modelling restraints enables generating three-dimensional models of native protein assemblies.

Due to its ability to directly reflect conformational changes and dynamics, H/D-X has been employed in many studies to investigate drug/ligand binding. A recent example is allosteric inhibition of MCL-1, a major pathologic factor in human cancer [66]. MCL-1 is an anti-apoptotic BCL-2 family protein binding to pro-apoptotic BCL-2 proteins through a surface groove thus preventing apoptosis of cancer cells. H/D-X and mass spectrometry helped elucidating an allosteric inhibition mechanism of MCL-1 by covalent modification through a small molecule [67]. Binding of MAIM1 (MCL-2 allosteric inhibition molecule 1) at an allosteric site causes conformational changes in MCL-2 and prevents binding of pro-apoptotic proteins [67].

2.4. Native ion mobility-mass spectrometry reveals shape and topology of protein assemblies

A versatile application in structural mass spectrometry is the analysis of intact protein complexes, a technique often called “native mass spectrometry”. It reveals protein interactions, stoichiometries, ligand binding and, in combination with IM, gives insights into the topology of the protein assemblies [68–70]. There are two pre-requisites, however: First, the protein assemblies have to be kept intact in the analysis buffer; contrary to the conventionally used organic solvents these are typically volatile, aqueous buffers. Second, the non-covalent protein interactions have to be maintained during analysis in the gas phase of the mass spectrometer. The latter is achieved by several instrument modifications mostly involving changes in the pressure to allow desolvation and gentle transfer of the macromolecules into the gas phase by collisional cooling [71]. In most cases, collisional induced dissociation (CID) is applied to strip subunits off intact protein assemblies, however, additional instrument modifications allow fragmentation by surface induced dissociation (SID) [72], infrared laser radiation [73], ultraviolet photodissociation (UVPD) [74,75] and electron transfer dissociation (ETD) [76] - all deliver complementary information on the structural arrangements. In this way, many intact protein assemblies including ribosomes [77], viruses [78,79], ATP synthases [80,81] were successfully studied and their stoichiometries and in some cases conformations were determined. For both, native MS and native IM-MS experiments, specialised software was developed to facilitate data analysis. These include *Massign* [82] and *Unidec* [83], which allow the determination of protein stoichiometries in heterogeneous assemblies and *PULSAR* [84] or *Amphitrite* [85] that enable the analysis of IM data including calculation of collisional cross sections. Further improvements in native MS also allowed the analysis of membrane protein complexes maintaining their interactions with lipids [86–88].

Many membrane proteins are drug targets and consequently are the key players during various disease. A recent study employed high resolution native mass spectrometry to unravel binding of HIV protease inhibitors, metal ions and detergents to human zinc protease ZMPSTE24. ZMPSTE24 is involved in processing of lamin A, which is one of the major components of nuclear lamina [89]. ZMPSTE24 deficiency causes progeroid syndromes, which mimic physiological aging [90]. Mass spectrometry enabled monitoring cleavage of prelamin A in real-time and, furthermore, showed that ZMPSTE24 protease activity was inhibited by HIV inhibitors revealing unknown side-effects of HIV drugs [91]. An example spectrum of HIV drug binding is shown in Fig. 1C.

3. Computational modelling techniques

3.1. Restraint-guided computational modelling

Computational strategies have emerged as invaluable tools for interpreting the structural and dynamic information extracted from complementary mass spectrometric techniques. The power of such methods lies in their ability to bring together information that otherwise will be reported independently. Typically, they involve two distinct but nonetheless related steps: a) sampling the protein or protein assembly configurations and, b) analysing the models generated. In the first step a search of protein configurational space is performed using established statistical averaging algorithms, i.e. Monte Carlo simulations and/or molecular dynamics (MD) such as replica-exchange MD. Depending on the structural information available, the proteins are represented by either low (e.g. coarse-grain, molecular envelopes or 2D images) or high-resolution structures (e.g. crystal or NMR structures, homology models), while in the cases of protein complex architectures hybrid representations can be employed [12]. These configurations are subjected to search for “good” structural models, guided by modelling restraints generated from the relevant experiments [92]. Such restraints can either be directly incorporated into the sampling algorithms by employing a scoring function [12,92] or used to filter structural models generated by computational methods [93,94]. Overall, the advantage of the first approach is that it can simultaneously integrate multiple types of restraints and as such to minimize the time required for downstream data analysis. The search will further be guided towards regions of the search space being more likely to be relevant. On the other hand, the second approach can yield structural models corresponding to the different restraints used without the need to repeat the sampling experiments. In either case the resulting model structures should match the input data and as such it would reflect the data used to generate them. For the purpose of this review we assume a modelling workflow that follows the first approach.

3.2. Ensemble analysis

Having generated the models matching the input experimental information, the second step of the integrative modelling workflow involves the data analysis and their subsequent interpretation. Ideally this should involve an iterative process where the structural models built are used to judge whether or not the experimental data are sufficient to guide us towards unambiguous solutions. It is therefore expected that in cases where the experimental information available are insufficient to yield credible model(s), additional experiments from the same or a different method are required. As such the data analysis and interpretation usually involves a model refinement step as well as computational algorithms to evaluate the distinctiveness of the solutions. The first allows enhancing the accuracy of the outputs by increasing the quality-of-fit between the models and the input data, while the second provides a measure of the precision of the method. In case of mass spectrometry-based integration, for which benchmark studies have shown that the accuracy of computational predictions is generally enhanced with increased amount of data used [12], it is often difficult to determine the level of accuracy one would expect from the various datasets. This is often depending on the nature of the system studied and the type of mass spectrometric methods used. For instance, relatively high resolution cross-linking MS methods have scored high with respect to the accuracy of 3D structural predictions in relevant benchmarking exercises [48,95]. On the other hand while lower resolution restraints such as subunit connectivity (native MS) [12], CCS (IM-MS) [96] and solvent accessibilities (covalent labelling MS) [9] exhibit limited predictability, their computational integration allowed more accurate predictions demonstrating the power of integrative modelling strategies. Overall a current challenge in the field is the development of integrative mass spectrometry-based workflows for generating

accurate three-dimensional models representing the optimal fit to the experiments.

3.3. Scoring structural models against multiple experimental data

Despite the advantages of the integrative modelling and docking strategies, they usually require careful handling of data uncertainties (e.g. false positives, errors) and of incomplete or heterogeneous information. To account for these, a Bayesian scoring function for guiding the sampling of protein configurational space can be followed. Such scoring function brings together the individual restraints by simultaneously accounting for a parameter that describes data uncertainties and/or heterogeneities. It is worth noting that careful consideration should be given when using a weighing function for the integration of restraints from diverse and heterogeneous sources. Weighting the individual terms is largely depending on the completeness, quality and amount of data obtained and may be subject of thorough analysis using known structures [48]. The scoring function finds the likelihood of a model structure given the input restraints derived from the various experiments [9,97]. Within this approach, a restraint can be expressed as a modelling feature (e.g. solvent accessibility, time-dependent uptake of deuterium, change in collision cross-section, ATD) quantified with respect to the data used to define it. A fully satisfied restraint can take a value of 0, while violations of all restraints would result to a value of 1

[96,98]. The most likely structural model will be the model that exhibits the least violation of restraints given the probability distribution.

3.4. Software tools for integrative modelling

Multi-scale software tools that enable integrative modelling of protein complex assembly and dynamics are the Integrative Modelling Platform [92] and HADDOCK [99], while docking algorithms based on geometric complementarity and energy considerations include for instance PatchDock [100] or Firedock [101]. Individual software modules and packages have so far used different mass spectrometric techniques to inform protein complex modelling. These include ion mobility [94,96], native MS [12], cross-linking MS [48,97,102,103], covalent labelling MS [9] and HDX [104] as well as the integration of MS-based techniques with EM [105] and NMR [93]. Computational modelling of heterogeneous datasets, including MS-based techniques, has been exemplified in particularly large and dynamic assemblies such as the nuclear pore complex [98], the eIF1:eIF3 translation initiation complex [106], the proteasome [48,103], amyloidogenic proteins [107–109] and the ATP synthase [81].

3.5. Modelling protein assembly dynamics

Despite all these noteworthy progresses, most of these tools are making predictions of static models without taking into account

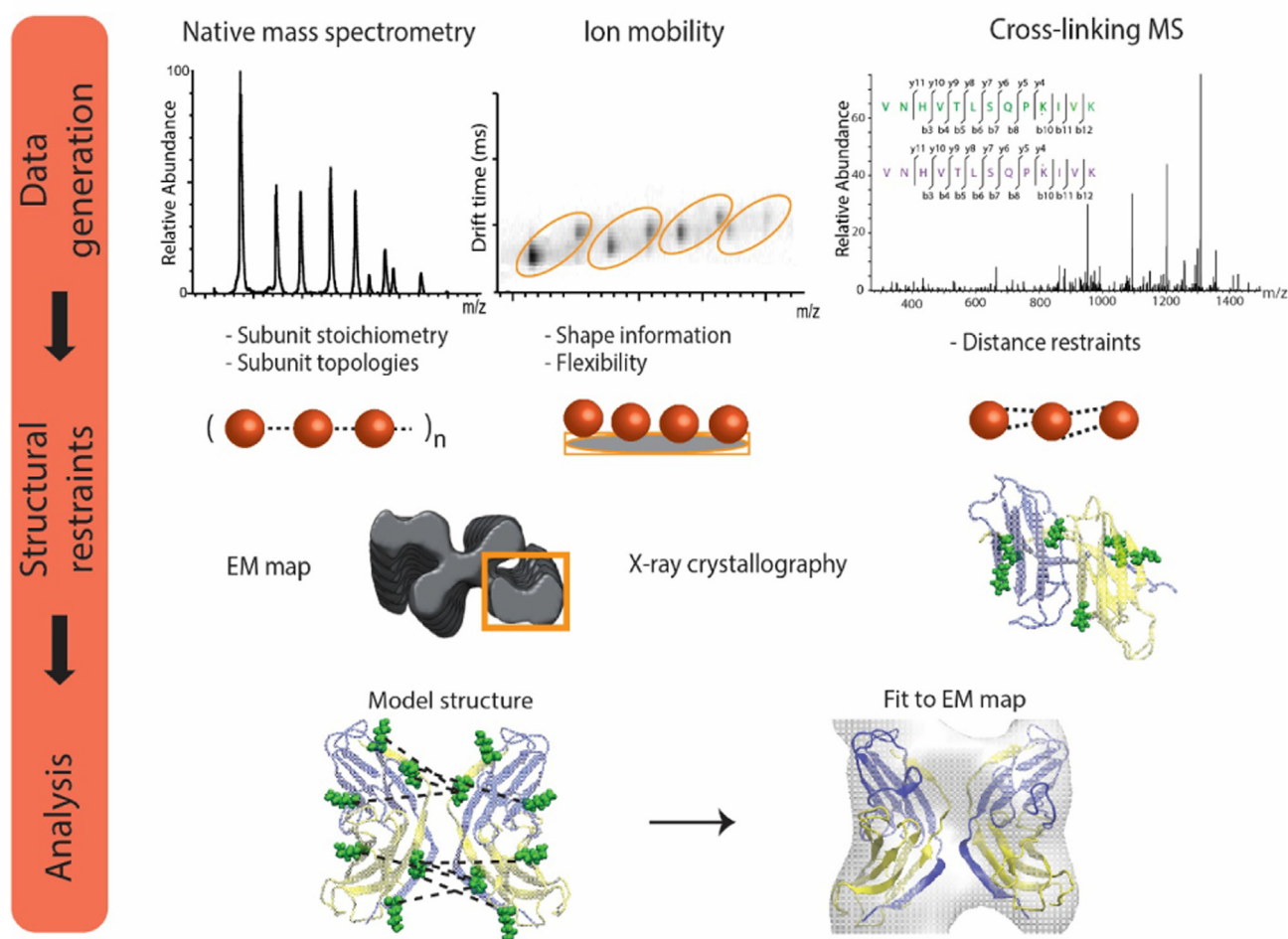


Fig. 2. Modelling workflow to generate structural models of $\beta 2$ -microglobulin oligomers. Various mass spectrometric techniques were employed to gain information on stoichiometry and topology (native MS), shape and flexibility (IM-MS) and protein-protein interactions (cross-linking). These results were converted into structural restraints which were integrated with structures from X-ray crystallography or EM. The generated models were analysed by fitting structural restraints into the available structures.

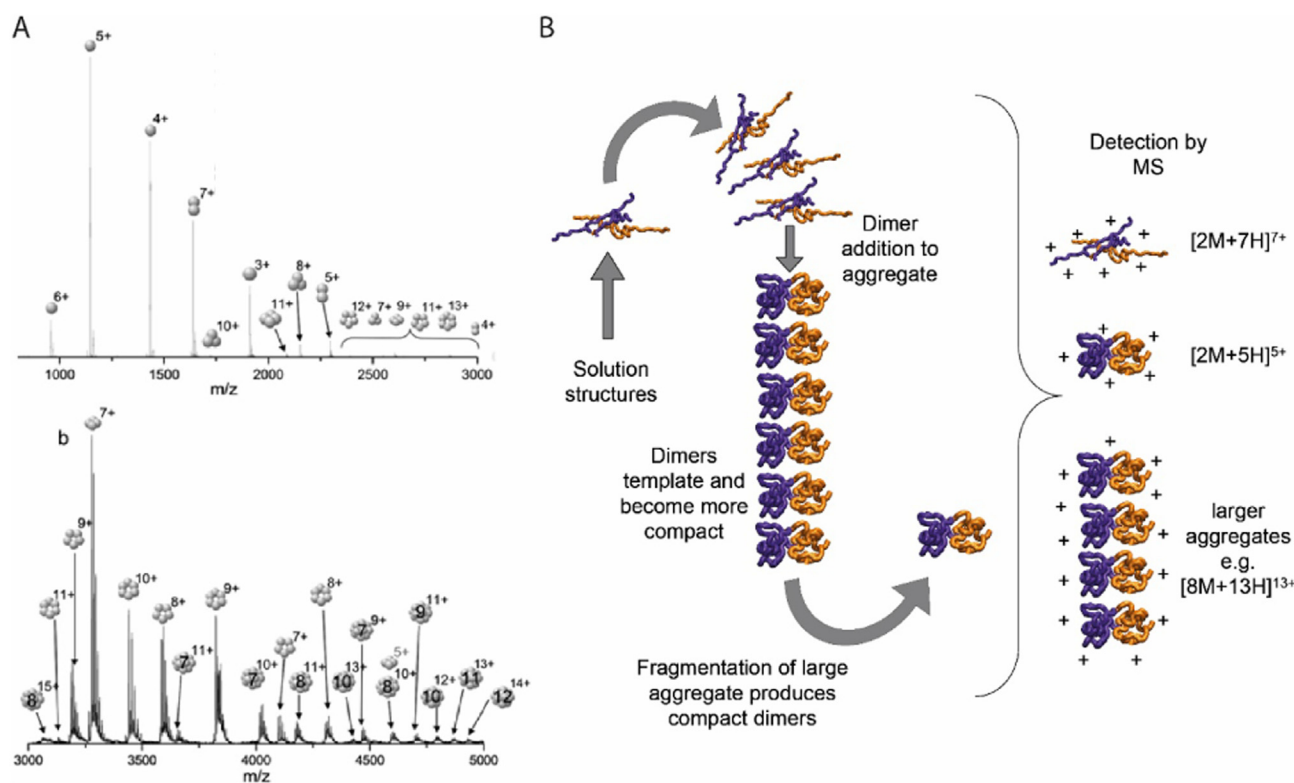


Fig. 3. Oligomerisation of insulin followed by IM-MS and molecular modelling. (A) Mass spectra of early oligomers reveal monomeric and dimeric insulin as the most abundant species (upper panel). Oligomers up to dodecamers were also observed (lower panel). (B) A novel assembly mechanism based on experimental data was proposed. The dimer forms the core unit in the higher aggregates. Adapted from [106].

small or large scale conformational changes of proteins and their complexes. This is more problematic for the case where slow conformational changes occur which require significant amount of computational resources. As such a current challenge in the integrative modelling strategies is the development of tools that can monitor protein dynamics at time-scales comparable to experiments [110]. This becomes pertinent with the recent developments in complementary MS-based techniques targeting protein dynamics such as HDX-MS [111,112]. The development of both the experimental MS-based methods and the computational algorithms for structural analysis holds great promise for enabling keen scientists to study the conformational states of large protein complexes, information that is often challenging to achieve using conventional strategies. Of note it is the potential of such tools to probe binding interfaces between protein and small molecules. As a consequence, it will allow researchers to generate testable hypotheses for the function of large protein complexes in the cell.

4. Integrating structural mass spectrometry and computational modelling - application examples

4.1. Early assembly mechanism for amyloidogenic β 2-microglobulin revealed by cross-linking and native mass spectrometry

A recent study highlighted the combination of mass spectrometry-based methods with integrative modelling and MD simulations for examining the mechanism of oligomer formation in β 2-microglobulin [107]. β 2-Microglobulin forms oligomers, which can aggregate into amyloid fibrils with serious clinical consequences (e.g. [113]). By combining native, ion mobility and chemical cross-linking MS through a modelling strategy, an early assembly pathway for this important protein was proposed. In particular, native MS allowed the identification of early oligomeric intermediates ranging from monomer to tetramer,

while IM-MS provided structural information that enabled shape restraints for interrogating computer-generated structures. This together with distance restraints from chemical cross-linking MS and by integrating with existing data from traditional structural techniques such as X-ray crystallography, NMR spectroscopy and electron microscopy, informed a modelling workflow (Fig. 2). In this workflow, a multi-scale modelling strategy was adopted comprising the use of a scoring function to bring together the data from various mass spectrometry-based sources and *in vacuo* molecular dynamics simulations to interrogate gas-phase model structures. This enabled the proposal of a step-wise assembly mechanism with domain-swapping for β 2-microglobulin that may precede fibril formation [107]. Overall, the hybrid mass spectrometry approach used in this study demonstrated the power of combined experimental and computational workflow for understanding clinically relevant protein complexes that are often challenging targets in structural biology.

4.2. Insulin fibrillation examined by ion mobility mass spectrometry and molecular modelling

Insulin is a peptide hormone which forms fibrils under amyloidogenic conditions. A recent study investigated early oligomers of insulin preceding fibril growth by combining IM-MS with molecular modelling [114]. While mass spectra revealed monomers and dimers being the most abundant species, higher oligomers, up to dodecamers, were also observed (Fig. 3A). Arrival time distributions of these oligomers obtained from IM-MS showed several conformations of these oligomers with the compact form in the highest abundance. IM-MS further allowed calculation of collision cross-sections of the oligomers.

The collisional cross-section and charge states obtained for dimeric insulin agreed well with the available crystal structure and the calculated net charge. The dimer was therefore split into two monomers (3 +

and 4+ species) which were subjected to molecular dynamics simulations. Various conformations of the two monomers were observed and employed in protein docking to produce a dimer which was then subjected to MD simulations. Docking of the compact monomers yielded conformations of the compact dimers, while docking the extended monomers yielded the extended dimers. This approach allowed generation of the conformational family observed experimentally. Investigating the contact interface and stability of the dimers revealed that hydrophobic interactions are fundamental in the oligomerisation process. Comparing the results from a previous study [115] with the obtained collisional cross sections and oligomeric species obtained suggest a novel assembly mechanism in which the dimeric species may be the core stacking species in larger aggregates (Fig. 3B).

5. Conclusions

We reviewed the emerging role of integrative mass spectrometry for predicting structural models of macromolecular assemblies. Overall, the future of structural mass spectrometry is hybrid and encapsulates the use of diverse methods for understanding the workings of biologically and medically relevant systems. We anticipate that the continuing development of mass spectrometry-based hybrid methodologies will contribute to consolidating the critical role of structural mass spectrometry in structural biology and will enable tackling challenges and limitations mainly associated with heterogeneity and dynamics currently frustrating individual approaches.

Conflict of interest

The authors declare that they have no conflicts of interests.

Acknowledgements

We thank Jasmin Corso, Henning Urlaub (both MPI for Biophysical Chemistry, Göttingen), Thomas Oellerich (Johann Wolfgang Goethe University Frankfurt), Rico Schwarz, Andrea Sinz (both Martin Luther University Halle-Wittenberg), Shahid Mehmood, Carol Robinson (both University of Oxford) and Perdita Barran (University of Manchester) for providing mass spectra and figures. A.P. acknowledges funding from the Wellcome Trust (109854/Z/15/z). C.S. acknowledges funding from the Federal Ministry of Education and Research (BMBF, ZIK programme; 03Z22HN22), the European Regional Development Fund (EFRE; ZS/2016/04/78115) and the MLU Halle-Wittenberg.

References

- [1] R. Aebersold, M. Mann, Mass-spectrometric exploration of proteome structure and function, *Nature* 537 (7620) (2016) 347–355.
- [2] M. Landreh, et al., A sliding selectivity scale for lipid binding to membrane proteins, *Curr. Opin. Struct. Biol.* 39 (2016) 54–60.
- [3] A. Leitner, et al., Crosslinking and mass spectrometry: an integrated technology to understand the structure and function of molecular machines, *Trends Biochem. Sci.* 41 (1) (2016) 20–32.
- [4] I. Liko, et al., Mass spectrometry guided structural biology, *Curr. Opin. Struct. Biol.* 40 (2016) 136–144.
- [5] C. Schmidt, K. Kramer, H. Urlaub, Investigation of protein-RNA interactions by mass spectrometry—techniques and applications, *J. Proteome* 75 (12) (2012) 3478–3494.
- [6] C. Schmidt, C.V. Robinson, Dynamic protein ligand interactions—insights from MS, *FEBS J.* 281 (8) (2014) 1950–1964.
- [7] A.D. Catherman, O.S. Skinner, N.L. Kelleher, Top down proteomics: facts and perspectives, *Biochem. Biophys. Res. Commun.* 445 (4) (2014) 683–693.
- [8] U. Leurs, U.H. Mistrarz, K.D. Rand, Getting to the core of protein pharmaceuticals—comprehensive structure analysis by mass spectrometry, *Eur. J. Pharm. Biopharm.* 93 (2015) 95–109.
- [9] C. Schmidt, et al., Surface accessibility and dynamics of macromolecular assemblies probed by covalent labeling mass spectrometry and integrative modeling, *Anal. Chem.* 89 (3) (2017) 1459–1468.
- [10] A. Sinz, et al., Chemical cross-linking and native mass spectrometry: a fruitful combination for structural biology, *Protein Sci.* 24 (8) (2015) 1193–1209.
- [11] K. Lasker, et al., Integrative structure modeling of macromolecular assemblies from proteomics data, *Mol. Cell. Proteomics* 9 (8) (2010) 1689–1702.
- [12] A. Politis, et al., Topological models of heteromeric protein assemblies from mass spectrometry: application to the yeast eIF3:eIF5 complex, *Chem. Biol.* 22 (1) (2015) 117–128.
- [13] A. Sali, T.L. Blundell, Comparative protein modelling by satisfaction of spatial restraints, *J. Mol. Biol.* 234 (3) (1993) 779–815.
- [14] M.M. Gromiha, K. Yugandhar, S. Jemimah, Protein–protein interactions: scoring schemes and binding affinity, *Curr. Opin. Struct. Biol.* 44 (2017) 31–38.
- [15] A. Schlössinger, et al., Structure-based discovery of prescription drugs that interact with the norepinephrine transporter, NET, *Proc. Natl. Acad. Sci. U. S. A.* 108 (38) (2011) 15810–15815.
- [16] A. Hospital, et al., Molecular dynamics simulations: advances and applications, *Adv. Appl. Bioinforma. Chem.* 8 (2015) 37–47.
- [17] M.R. Wilkins, et al., Progress with proteome projects: why all proteins expressed by a genome should be identified and how to do it, *Biotechnol. Genet. Eng. Rev.* 13 (1996) 19–50.
- [18] C.H. Ahrens, et al., Generating and navigating proteome maps using mass spectrometry, *Nat. Rev. Mol. Cell Biol.* 11 (11) (2010) 789–801.
- [19] X. Han, A. Aslanian, J.R. Yates 3rd, Mass spectrometry for proteomics, *Curr. Opin. Chem. Biol.* 12 (5) (2008) 483–490.
- [20] M. Mann, R.C. Hendrickson, A. Pandey, Analysis of proteins and proteomes by mass spectrometry, *Annu. Rev. Biochem.* 70 (2001) 437–473.
- [21] T.C. Walther, M. Mann, Mass spectrometry-based proteomics in cell biology, *J. Cell Biol.* 190 (4) (2010) 491–500.
- [22] H. Steen, M. Mann, The ABC's (and XYZ's) of peptide sequencing, *Nat. Rev. Mol. Cell Biol.* 5 (9) (2004) 699–711.
- [23] S. Camerini, P. Mauri, The role of protein and peptide separation before mass spectrometry analysis in clinical proteomics, *J. Chromatogr. A* 1381 (2015) 1–12.
- [24] A. Motoyama, J.R. Yates 3rd, Multidimensional LC separations in shotgun proteomics, *Anal. Chem.* 80 (19) (2008) 7187–7193.
- [25] T. Stasyk, L.A. Huber, Zooming in: fractionation strategies in proteomics, *Proteomics* 4 (12) (2004) 3704–3716.
- [26] J.V. Olsen, M. Mann, Status of large-scale analysis of post-translational modifications by mass spectrometry, *Mol. Cell. Proteomics* 12 (12) (2013) 3444–3452.
- [27] Y. Zhao, O.N. Jensen, Modification-specific proteomics: strategies for characterization of post-translational modifications using enrichment techniques, *Proteomics* 9 (20) (2009) 4632–4641.
- [28] M. Bantscheff, et al., Quantitative mass spectrometry in proteomics: critical review update from 2007 to the present, *Anal. Bioanal. Chem.* 404 (4) (2012) 939–965.
- [29] M. Bantscheff, et al., Quantitative mass spectrometry in proteomics: a critical review, *Anal. Bioanal. Chem.* 389 (4) (2007) 1017–1031.
- [30] S. Beck, et al., The impact II, a very high-resolution quadrupole time-of-flight instrument (QTOF) for deep shotgun proteomics, *Mol. Cell. Proteomics* 14 (7) (2015) 2014–2029.
- [31] M. Mann, et al., The coming age of complete, accurate, and ubiquitous proteomes, *Mol. Cell* 49 (4) (2013) 583–590.
- [32] L.M. de Godoy, et al., Status of complete proteome analysis by mass spectrometry: SILAC labeled yeast as a model system, *Genome Biol.* 7 (6) (2006) R50.
- [33] A.C. Gavin, et al., Proteome survey reveals modularity of the yeast cell machinery, *Nature* 440 (7084) (2006) 631–636.
- [34] N.J. Krogan, et al., Global landscape of protein complexes in the yeast *Saccharomyces cerevisiae*, *Nature* 440 (7084) (2006) 637–643.
- [35] E. Brunner, et al., A high-quality catalog of the *Drosophila melanogaster* proteome, *Nat. Biotechnol.* 25 (5) (2007) 576–583.
- [36] K.G. Guruharsha, et al., A protein complex network of *Drosophila melanogaster*, *Cell* 147 (3) (2011) 690–703.
- [37] R.M. Ewing, et al., Large-scale mapping of human protein-protein interactions by mass spectrometry, *Mol. Syst. Biol.* 3 (2007) 89.
- [38] M.S. Kim, et al., A draft map of the human proteome, *Nature* 509 (7502) (2014) 575–581.
- [39] M. Wilhelm, et al., Mass-spectrometry-based draft of the human proteome, *Nature* 509 (7502) (2014) 582–587.
- [40] Y. Liu, et al., Mass spectrometric protein maps for biomarker discovery and clinical research, *Expert. Rev. Mol. Diagn.* 13 (8) (2013) 811–825.
- [41] M. Palmblad, A. Tiss, R. Cramer, Mass spectrometry in clinical proteomics - from the present to the future, *Proteomics Clin. Appl.* 3 (1) (2009) 6–17.
- [42] C.E. Parker, et al., Mass-spectrometry-based clinical proteomics—a review and perspective, *Analyst* 135 (8) (2010) 1830–1838.
- [43] C. Choudhary, M. Mann, Decoding signalling networks by mass spectrometry-based proteomics, *Nat. Rev. Mol. Cell Biol.* 11 (6) (2010) 427–439.
- [44] J. Corso, et al., Elucidation of tonic and activated B-cell receptor signaling in Burkitt's lymphoma provides insights into regulation of cell survival, *Proc. Natl. Acad. Sci. U. S. A.* 113 (20) (2016) 5688–5693.
- [45] T. Kurosaki, Regulation of BCR signaling, *Mol. Immunol.* 48 (11) (2011) 1287–1291.
- [46] J. Rappsilber, The beginning of a beautiful friendship: cross-linking/mass spectrometry and modelling of proteins and multi-protein complexes, *J. Struct. Biol.* 173 (3) (2011) 530–540.
- [47] A. Sinz, Chemical cross-linking and mass spectrometry to map three-dimensional protein structures and protein-protein interactions, *Mass Spectrom. Rev.* 25 (4) (2006) 663–682.
- [48] A. Politis, et al., A mass spectrometry-based hybrid method for structural modeling of protein complexes, *Nat. Methods* 11 (4) (2014) 403–406.

- [49] M. Schneider, et al., Blind testing of cross-linking/mass spectrometry hybrid methods in CASP11, *Proteins* 84 (Suppl. 1) (2016) 152–163.
- [50] O. Rinner, et al., Identification of cross-linked peptides from large sequence databases, *Nat. Methods* 5 (4) (2008) 315–318.
- [51] H. Xu, et al., Database search algorithm for identification of intact cross-links in proteins and peptides using tandem mass spectrometry, *J. Proteome Res.* 9 (7) (2010) 3384–3393.
- [52] M. Gotze, et al., StavroX—a software for analyzing crosslinked products in protein interaction studies, *J. Am. Soc. Mass Spectrom.* 23 (1) (2012) 76–87.
- [53] B. Yang, et al., Identification of cross-linked peptides from complex samples, *Nat. Methods* 9 (9) (2012) 904–906.
- [54] D.E. Agafonov, et al., Molecular architecture of the human U4/U6.U5 tri-snRNP, *Science* 351 (6280) (2016) 1416–1420.
- [55] R.G. Efremov, et al., Architecture and conformational switch mechanism of the ryanodine receptor, *Nature* 517 (7532) (2015) 39–43.
- [56] B.J. Greber, et al., Ribosome. The complete structure of the 55S mammalian mitochondrial ribosome, *Science* 348 (6232) (2015) 303–308.
- [57] R. Rauhut, et al., Molecular architecture of the *Saccharomyces cerevisiae* activated spliceosome, *Science* 353 (6306) (2016) 1399–1405.
- [58] B. Desvergne, W. Wahli, Peroxisome proliferator-activated receptors: nuclear control of metabolism, *Endocr. Rev.* 20 (5) (1999) 649–688.
- [59] V. Souza-Mello, Peroxisome proliferator-activated receptors as targets to treat non-alcoholic fatty liver disease, *World J. Hepatol.* 7 (8) (2015) 1012–1019.
- [60] B. Staels, J.C. Fruchart, Therapeutic roles of peroxisome proliferator-activated receptor agonists, *Diabetes* 54 (8) (2005) 2460–2470.
- [61] R. Schwarz, et al., Monitoring solution structures of peroxisome proliferator-activated receptor beta/delta upon ligand binding, *PLoS One* 11 (3) (2016), e0151412.
- [62] S.R. Marcisin, J.R. Engen, Hydrogen exchange mass spectrometry: what is it and what can it tell us? *Anal. Bioanal. Chem.* 397 (3) (2010) 967–972.
- [63] Z. Zhang, D.L. Smith, Determination of amide hydrogen exchange by mass spectrometry: a new tool for protein structure elucidation, *Protein Sci.* 2 (4) (1993) 522–531.
- [64] A. Miranker, et al., Detection of transient protein folding populations by mass spectrometry, *Science* 262 (5135) (1993) 896–900.
- [65] C.V. Robinson, et al., Probing the nature of noncovalent interactions by mass spectrometry. A study of protein-CoA ligand binding and assembly, *J. Am. Chem. Soc.* 118 (1996) 8646–8653.
- [66] R. Beroukhi, et al., The landscape of somatic copy-number alteration across human cancers, *Nature* 463 (7283) (2010) 899–905.
- [67] S. Lee, et al., Allosteric inhibition of antiapoptotic MCL-1, *Nat. Struct. Mol. Biol.* 23 (6) (2016) 600–607.
- [68] A.J. Heck, Native mass spectrometry: a bridge between interactomics and structural biology, *Nat. Methods* 5 (11) (2008) 927–933.
- [69] S. Mehmood, T.M. Allison, C.V. Robinson, Mass spectrometry of protein complexes: from origins to applications, *Annu. Rev. Phys. Chem.* 66 (2015) 453–474.
- [70] M. Sharon, C.V. Robinson, The role of mass spectrometry in structure elucidation of dynamic protein complexes, *Annu. Rev. Biochem.* 76 (2007) 167–193.
- [71] F. Sobott, et al., A tandem mass spectrometer for improved transmission and analysis of large macromolecular assemblies, *Anal. Chem.* 74 (6) (2002) 1402–1407.
- [72] M. Zhou, S. Dagan, V.H. Wysocki, Impact of charge state on gas-phase behaviors of noncovalent protein complexes in collision induced dissociation and surface induced dissociation, *Analyst* 138 (5) (2013) 1353–1362.
- [73] V.A. Mikhailov, et al., Infrared laser activation of soluble and membrane protein assemblies in the gas phase, *Anal. Chem.* 88 (14) (2016) 7060–7067.
- [74] D.D. Holden, J.S. Brodbelt, Ultraviolet photodissociation of native proteins following proton transfer reactions in the gas phase, *Anal. Chem.* 88 (24) (2016) 12354–12362.
- [75] S. Tamara, et al., Symmetry of charge partitioning in collisional and UV photon-induced dissociation of protein assemblies, *J. Am. Chem. Soc.* 138 (34) (2016) 10860–10868.
- [76] F. Lermite, F. Sobott, Electron transfer dissociation provides higher-order structural information of native and partially unfolded protein complexes, *Proteomics* 15 (16) (2015) 2813–2822.
- [77] A.A. Rostom, et al., Detection and selective dissociation of intact ribosomes in a mass spectrometer, *Proc. Natl. Acad. Sci. U. S. A.* 97 (10) (2000) 5185–5190.
- [78] J. Snijder, et al., Defining the stoichiometry and cargo load of viral and bacterial nanoparticles by Orbitrap mass spectrometry, *J. Am. Chem. Soc.* 136 (20) (2014) 7295–7299.
- [79] C. Uetrecht, et al., Stability and shape of hepatitis B virus capsids in vacuo, *Angew. Chem. Int. Ed. Engl.* 47 (33) (2008) 6247–6251.
- [80] C. Schmidt, et al., Comparative cross-linking and mass spectrometry of an intact F-type ATPase suggest a role for phosphorylation, *Nat. Commun.* 4 (2013) 1985.
- [81] M. Zhou, et al., Mass spectrometry of intact V-type ATPases reveals bound lipids and the effects of nucleotide binding, *Science* 334 (6054) (2011) 380–385.
- [82] N. Morgner, C.V. Robinson, Massign: an assignment strategy for maximizing information from the mass spectra of heterogeneous protein assemblies, *Anal. Chem.* 84 (6) (2012) 2939–2948.
- [83] M.T. Marty, et al., Bayesian deconvolution of mass and ion mobility spectra: from binary interactions to polydisperse ensembles, *Anal. Chem.* 87 (8) (2015) 4370–4376.
- [84] T.M. Allison, et al., Quantifying the stabilizing effects of protein-ligand interactions in the gas phase, *Nat. Commun.* 6 (2015) 8551.
- [85] G.N. Sivalingam, et al., Amphitrite: a program for processing travelling wave ion mobility mass spectrometry data, *Int. J. Mass Spectrom.* 345–347 (2013) 54–62.
- [86] C. Bechara, et al., A subset of annular lipids is linked to the flippase activity of an ABC transporter, *Nat. Chem.* 7 (3) (2015) 255–262.
- [87] K. Gupta, et al., The role of interfacial lipids in stabilizing membrane protein oligomers, *Nature* 541 (7637) (2017) 421–424.
- [88] A. Laganowsky, et al., Membrane proteins bind lipids selectively to modulate their structure and function, *Nature* 510 (7503) (2014) 172–175.
- [89] R.D. Goldman, et al., Nuclear lamins: building blocks of nuclear architecture, *Genes Dev.* 16 (5) (2002) 533–547.
- [90] M. Eriksson, et al., Recurrent de novo point mutations in lamin A cause Hutchinson-Gilford progeria syndrome, *Nature* 423 (6937) (2003) 293–298.
- [91] S. Mehmood, et al., Mass spectrometry captures off-target drug binding and provides mechanistic insights into the human metalloprotease ZMPSTE24, *Nat. Chem.* 8 (12) (2016) 1152–1158.
- [92] D. Russel, et al., Putting the pieces together: integrative modeling platform software for structure determination of macromolecular assemblies, *PLoS Biol.* 10 (1) (2012), e1001244.
- [93] A.J. Baldwin, et al., The polydispersity of alphaB-crystallin is rationalized by an interconverting polyhedral architecture, *Structure* 19 (12) (2011) 1855–1863.
- [94] A. Politis, et al., Integrating ion mobility mass spectrometry with molecular modelling to determine the architecture of multiprotein complexes, *PLoS One* 5 (8) (2010), e12080.
- [95] Z. Chen, et al., Quantitative cross-linking/mass spectrometry reveals subtle protein conformational changes, *Wellcome Open Res.* 1 (2016) 5.
- [96] Z. Hall, A. Politis, C.V. Robinson, Structural modeling of heteromeric protein complexes from disassembly pathways and ion mobility-mass spectrometry, *Structure* 20 (9) (2012) 1596–1609.
- [97] Y. Shi, et al., A strategy for dissecting the architectures of native macromolecular assemblies, *Nat. Methods* 12 (12) (2015) 1135–1138.
- [98] F. Alber, et al., Integrating diverse data for structure determination of macromolecular assemblies, *Annu. Rev. Biochem.* 77 (2008) 443–477.
- [99] C. Dominguez, R. Boelens, A.M. Bonvin, HADDOCK: a protein-protein docking approach based on biochemical or biophysical information, *J. Am. Chem. Soc.* 125 (7) (2003) 1731–1737.
- [100] D. Schneidman-Duhovny, et al., PatchDock and SymmDock: servers for rigid and symmetric docking, *Nucleic Acids Res.* 33 (2005) W363–W367 (Web Server issue).
- [101] N. Andrusier, R. Nussinov, H.J. Wolfson, FireDock: fast interaction refinement in molecular docking, *Proteins* 69 (1) (2007) 139–159.
- [102] M. Ferber, et al., Automated structure modeling of large protein assemblies using crosslinks as distance restraints, *Nat. Methods* 13 (6) (2016) 515–520.
- [103] K. Lasker, et al., Molecular architecture of the 26S proteasome holocomplex determined by an integrative approach, *Proc. Natl. Acad. Sci. U. S. A.* 109 (5) (2012) 1380–1387.
- [104] D.J. Saltzberg, et al., A residue resolved Bayesian approach to quantitative interpretation of hydrogen deuterium exchange from mass spectrometry: application to characterizing protein-ligand interactions, *J. Phys. Chem. B* 121 (15) (2017) 3493–3501.
- [105] M.T. Degiacomi, J.L. Benesch, EM intersectionIM: software for relating ion mobility mass spectrometry and electron microscopy data, *Analyst* 141 (1) (2016) 70–75.
- [106] J.P. Erzberger, et al., Molecular architecture of the 40S ribosomal translation initiation complex, *Cell* 158 (5) (2014) 1123–1135.
- [107] Z. Hall, C. Schmidt, A. Politis, Uncovering the early assembly mechanism for amyloidogenic beta2-microglobulin using cross-linking and native mass spectrometry, *J. Biol. Chem.* 291 (9) (2016) 4626–4637.
- [108] J. Seo, et al., An infrared spectroscopy approach to follow β -sheet formation in peptide amyloid assemblies, *Nat. Chem.* 9 (1) (2017) 39–44.
- [109] D.P. Smith, S.E. Radford, A.E. Ashcroft, Elongated oligomers in beta2-microglobulin amyloid assembly revealed by ion mobility spectrometry-mass spectrometry, *Proc. Natl. Acad. Sci. U. S. A.* 107 (15) (2010) 6794–6798.
- [110] Z. Ahdash, E. Pyle, A. Politis, Hybrid mass spectrometry: towards characterization of protein conformational states, *Trends Biochem. Sci.* 41 (8) (2016) 650–653.
- [111] V. Katta, B.T. Chait, Conformational changes in proteins probed by hydrogen-exchange electrospray-ionization mass spectrometry, *Rapid Commun. Mass Spectrom.* 5 (4) (1991) 214–217.
- [112] T.E. Wales, J.R. Engen, Hydrogen exchange mass spectrometry for the analysis of protein dynamics, *Mass Spectrom. Rev.* 25 (1) (2006) 158–170.
- [113] J. Floege, et al., Clearance and synthesis rates of beta 2-microglobulin in patients undergoing hemodialysis and in normal subjects, *J. Lab. Clin. Med.* 118 (2) (1991) 153–165.
- [114] H. Cole, et al., Early stages of insulin fibrillogenesis examined with ion mobility mass spectrometry and molecular modelling, *Analyst* 140 (20) (2015) 7000–7011.
- [115] J.L. Jimenez, et al., The protofibril structure of insulin amyloid fibrils, *Proc. Natl. Acad. Sci. U. S. A.* 99 (14) (2002) 9196–9201.